

SEQUENCE LISTING

(1) GENERAL INFORMATION:

- (i) APPLICANT:
- (A) NAME: E. I. DUPONT DE NEMOURS AND COMPANY
 - (B) STREET: 1007 MARKET STREET
 - (C) CITY: WILMINGTON
 - (D) STATE: DELAWARE
 - (E) COUNTRY: U.S.A.
 - (F) POSTAL CODE (ZIP): 19898
 - (G) TELEPHONE: 302-892-8112
 - (H) TELEFAX: 302-773-0164
- (i) APPLICANT:
- (A) NAME: GENENCOR INTERNATIONAL, INC.
 - (B) STREET: 4 CAMBRIDGE PLACE
1870 SOUTH WINTON ROAD
 - (C) CITY: ROCHESTER
 - (D) STATE: NEW YORK
 - (E) COUNTRY: U.S.A.
 - (F) POSTAL CODE (ZIP): 14618
 - (G) TELEPHONE:
 - (H) TELEFAX:
- (ii) TITLE OF INVENTION: BIOCONVERSION OF A FERMENTABLE
CARBON SOURCE TO 1,3-PROPANE-
DIOL BY A SINGLE MICROORGANISM
- (iii) NUMBER OF SEQUENCES: 46
- (iv) COMPUTER READABLE FORM:
- (A) MEDIUM TYPE: 3.50 INCH DISKETTE
 - (B) COMPUTER: IBM
 - (C) OPERATING SYSTEM: MICROSOFT WINDOWS 3.1
 - (D) SOFTWARE: MICROSOFT WORD 6.0
- (v) CURRENT APPLICATION DATA:
- (A) APPLICATION NUMBER:
 - (B) FILING DATE:
 - (C) CLASSIFICATION:
- (vi) PRIOR APPLICATION DATA:
- (A) APPLICATION NUMBER: 08/440,293
 - (B) FILING DATE: MAY 12, 1995
- (vii) ATTORNEY/AGENT INFORMATION:
- (A) NAME: LINDA AXAMETHY FLOYD
 - (B) REGISTRATION NUMBER: 33,692
 - (C) REFERENCE/DOCKET NUMBER: CR-9715-B

(2) INFORMATION FOR SEQ ID NO:1:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 12145 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

| | |
|--|------|
| GTCGACCACC ACGGTGGTGA CTTTAATGCC GCTCTCATGC AGCAGCTCGG TGGCGGTCTC | 60 |
| AAAATTCAGG ATGTCGCCGG TATAGTTTTT GATAATCAGC AAGACGCCTT CGCCGCCGTC | 120 |
| AATTTGCATC GCGCATTCAA ACATTTTGTC CGGCGTCGGC GAGGTGAATA TTTCCCCCGG | 180 |
| ACAGGCGCCG GAGAGCATGC CCTGGCCGAT ATAGCCGCAG TGCATCGGTT CATGTCCGCT | 240 |
| GCCGCCGCCG GAGAGCAGGG CCACCTTGCC AGCCACCGGC GCGTCGGTGC GGGTCACATA | 300 |
| CAGCGGGTCC TGATGCAGGG TCAGCTGCGG ATGGGCTTTA GCCAGCCCTT GTAATTGTTC | 360 |
| ATTCAGTACA TCTTCAACAC GGTAAATCAG CTTTTTCATT ATTCAGTGCT CCGTTGGAGA | 420 |
| AGGTTTCGATG CCGCCTCTCT GCTGGCGGAG GCGGTCATCG CGTAGGGGTA TCGTCTGACG | 480 |
| GTGGAGCGTG CCTGGCGATA TGATGATTCT GGCTGAGCGG ACGAAAAAAA GAATGCCCCG | 540 |
| ACGATCGGGT TTCATTACGA AACATTGCTT CCTGATTTTG TTTCTTTATG GAACGTTTTT | 600 |
| GCTGAGGATA TGGTGAAAAT GCGAGCTGGC GCGCTTTTTT TCTTCTGCCA TAAGCGGCGG | 660 |
| TCAGGATAGC CGGCGAAGCG GGTGGGAAAA AATTTTTTGC TGATTTTCTG CCGACTGCGG | 720 |
| GAGAAAAGGC GGTCAAACAC GGAGGATTGT AAGGGCATTG TCGGCAAGAG GAGCGGATCG | 780 |
| GGATCGCAAT CCTGACAGAG ACTAGGGTTT TTTGTTCCAA TATGGAACGT AAAAAATTAA | 840 |
| CCTGTGTTTC ATATCAGAAC AAAAAGGCGA AAGATTTTTT TGTTCCCTGC CGGCCCTACA | 900 |
| GTGATCGCAC TGCTCCGGTA CGCTCCGTTC AGGCCGCGCT TCACTGGCCG GCGCGGATAA | 960 |
| CGCCAGGGCT CATCATGTCT ACATGCGCAC TTATTTGAGG GTGAAAGGAA TGCTAAAAGT | 1020 |
| TATTCAATCT CCAGCCAAAT ATCTTCAGGG TCCTGATGCT GCTGTTCTGT TCGGTCAATA | 1080 |
| TGCCAAAAAC CTGGCGGAGA GCTTCTTCGT CATCGCTGAC GATTTTCGTAA TGAAGCTGGC | 1140 |
| GGGAGAGAAA GTGGTGAATG GCCTGCAGAG CCACGATATT CGCTGCCATG CGGAACGGTT | 1200 |
| TAACGGCGAA TGCAGCCATG CGGAAATCAA CCGTCTGATG GCGATTTTGC AAAAACAGGG | 1260 |
| CTGCCGCGGC GTGGTCGGGA TCGGCGGTGG TAAAACCCTC GATACCGCGA AGGCGATCGG | 1320 |
| TTACTACCAG AAGCTGCCGG TGGTGGTGAT CCCGACCATC GCCTCGACCG ATGCGCCAAC | 1380 |

| | | | | | | |
|-------------|-------------|-------------|------------|------------|-------------|------|
| CAGCGCGCTG | TCGGTGATCT | ACACCGAAGC | GGGCGAGTTT | GAAGAGTATC | TGATCTATCC | 1440 |
| GAAAAACCCG | GATATGGTGG | TGATGGACAC | GGCGATTATC | GCCAAAGCGC | CGGTACGCCT | 1500 |
| GCTGGTCTCC | GGCATGGGCG | ATGCGCTCTC | CACCTGGTTC | GAGGCCAAAG | CTTGCTACGA | 1560 |
| TGCGCGCGCC | ACCAGCATGG | CCGGAGGACA | GTCCACCGAG | GCGGCGCTGA | GCCTCGCCCG | 1620 |
| CCTGTGCTAT | GATACGCTGC | TGGCGGAGGG | CGAAAAGGCC | CGTCTGGCGG | CGCAGGCCCG | 1680 |
| GGTAGTGACC | GAAGCGCTGG | AGCGCATCAT | CGAGGCGAAC | ACTTACCTCA | GCGGCATTGG | 1740 |
| CTTTGAAAGC | AGTGGCCTGG | CCGCTGCCCC | TGCAATCCAC | AACGGTTTCA | CCATTCTTGA | 1800 |
| AGAGTGCCAT | CACCTGTATC | ACGGTGAGAA | AGTGGCCTTC | GGTACCCTGG | CGCAGCTGGT | 1860 |
| GCTGCAGAAC | AGCCCGATGG | ACGAGATTGA | AACGGTGCAG | GGCTTCTGCC | AGCGCGTCGG | 1920 |
| CCTGCCGGTG | ACGCTCGCGC | AGATGGGCGT | CAAAGAGGGG | ATCGACGAGA | AAATCGCCGC | 1980 |
| GGTGGCGAAA | GCTACCTGCG | CGGAAGGGGA | AACCATCCAT | AATATGCCGT | TTGCCGTGAC | 2040 |
| CCCGGAGAGC | GTCCATGCCG | CTATCCTCAC | CGCCGATCTG | TTAGGCCAGC | AGTGGCTGGC | 2100 |
| GCGTTAATTC | GCGGTGGCTA | AACCGCTGGC | CCAGGTCAGC | GGTTTTTCTT | TCTCCCCTCC | 2160 |
| GGCAGTCGCT | GCCGGAGGGG | TTCTCTATGG | TACAACGCGG | AAAAGGATAT | GACTGTTCAG | 2220 |
| ACTCAGGATA | CCGGGAAGGC | GGTCTCTTCC | GTCATTGCCC | AGTCATGGCA | CCGCTGCAGC | 2280 |
| AAGTTTATGC | AGCGCGAAAC | CTGGCAAACG | CCGCACCAGG | CCCAGGGCCT | GACCTTCGAC | 2340 |
| TCCATCTGTC | GGCGTAA AAC | CGCGCTGCTC | ACCATCGGCC | AGGCGGCGCT | GGAAGACGCC | 2400 |
| TGGGAGTTTA | TGGACGGCCG | CCCCTGCGCG | CTGTTTATTC | TTGATGAGTC | CGCCTGCATC | 2460 |
| CTGAGCCGTT | GCGGCGAGCC | GCAAACCCCTG | GCCCAGCTGG | CTGCCCTGGG | ATTTTCGCGAC | 2520 |
| GGCAGCTATT | GTGCGGAGAG | CATTATCGGC | ACCTGCGCGC | TGTCGCTGGC | CGCGATGCAG | 2580 |
| GGCCAGCCGA | TCAACACCGC | CGGCGATCGG | CATTTTAAGC | AGGCGCTACA | GCCATGGAGT | 2640 |
| TTTTGCTCGA | CGCCGGTGTT | TGATAACCAC | GGGCGGCTGT | TCGGCTCTAT | CTCGCTTTGC | 2700 |
| TGTCTGGTCG | AGCACCAGTC | CAGCGCCGAC | CTCTCCCTGA | CGCTGGCCAT | CGCCCGCGAG | 2760 |
| GTGGGTA ACT | CCCTGCTTAC | CGACAGCCTG | CTGGCGGAAT | CCAACCGTCA | CCTCAATCAG | 2820 |
| ATGTACGGCC | TGCTGGAGAG | CATGGACGAT | GGGGTGATGG | CGTGGAACGA | ACAGGGCGTG | 2880 |
| CTGCAGTTTC | TCAATGTTCA | GGCGGCGAGA | CTGCTGCATC | TTGATGCTCA | GGCCAGCCAG | 2940 |
| GGGAAAAATA | TCGCCGATCT | GGTGACCCTC | CCGGCGCTGC | TGCGCCGCGC | CATCAAACAC | 3000 |
| GCCCGCGGCC | TGAATCACGT | CGAAGTCACC | TTTGAAAGTC | AGCATCAGTT | TGTCGATGCG | 3060 |
| GTGATCACCT | TAAAACCGAT | TGTCGAGGCG | CAAGGCAACA | GTTTTATTCT | GCTGCTGCAT | 3120 |

09575638.052200

| | | | | | | |
|------------|------------|------------|-------------|------------|-------------|------|
| CCGGTGGAGC | AGATGCGGCA | GCTGATGACC | AGCCAGCTCG | GTAAAGTCAG | CCACACCTTT | 3180 |
| GAGCAGATGT | CTGCCGACGA | TCCGGAAACC | CGACGCCTGA | TCCACTTTGG | CCGCCAGGCG | 3240 |
| GCGCGCGGCG | GCTTCCCGGT | GCTACTGTGC | GGCGAAGAGG | GGGTCGGGAA | AGAGCTGCTG | 3300 |
| AGCCAGGCTA | TTCACAATGA | AAGCGAACGG | GCGGGCGGCC | CCTACATCTC | CGTCAACTGC | 3360 |
| CAGCTATATG | CCGACAGCGT | GCTGGGCCAG | GACTTTATGG | GCAGCGCCCC | TACCGACGAT | 3420 |
| GAAAATGGTC | GCCTGAGCCG | CCTTGAGCTG | GCCAACGGCG | GCACCCTGTT | TCTGGAAAAG | 3480 |
| ATCGAGTATC | TGGCGCCGGA | GCTGCAGTCG | GCTCTGCTGC | AGGTGATTAA | GCAGGGCGTG | 3540 |
| CTCACCCGCC | TCGACGCCCC | GCGCCTGATC | CCGGTGGATG | TGAAGGTGAT | TGCCACCACC | 3600 |
| ACCGTCGATC | TGGCCAATCT | GGTGGAACAG | AACCGCTTTA | GCCGCCAGCT | GTACTATGCG | 3660 |
| CTGCACTCCT | TTGAGATCGT | CATCCCGCCG | CTGCGCGCCC | GACGCAACAG | TATTCCGTCG | 3720 |
| CTGGTGCATA | ACCGGTTGAA | GAGCCTGGAG | AAGCGTTTCT | CTTCGCGACT | GAAAAGTGGAC | 3780 |
| GATGACGCGC | TGGCACAGCT | GGTGGCCTAC | TCGTGGCCGG | GGAATGATTT | TGAGCTCAAC | 3840 |
| AGCGTCATTG | AGAATATCGC | CATCAGCAGC | GACAACGGCC | ACATTGCGCT | GAGTAATCTG | 3900 |
| CCGGAATATC | TCTTTTCCGA | GCGGCCGGGC | GGGGATAGCG | CGTCATCGCT | GCTGCCGGCC | 3960 |
| AGCCTGACTT | TTAGCGCCAT | CGAAAAGGAA | GCTATTATTC | ACGCCGCCCC | GGTGACCAGC | 4020 |
| GGGCGGGTGC | AGGAGATGTC | GCAGCTGCTC | AATATCGGCC | GCACCACCCT | GTGGCGCAAA | 4080 |
| ATGAAGCAGT | ACGATATTGA | CGCCAGCCAG | TTCAAGCGCA | AGCATCAGGC | CTAGTCTCTT | 4140 |
| CGATTGCGCG | CATGGAGAAC | AGGGCATCCG | ACAGGCGATT | GCTGTAGCGT | TTGAGCGCGT | 4200 |
| CGCGCAGCGG | ATGCGCGCGG | TCCATGGCCG | TCAGCAGGCG | TTGAGCCGA | CGGGACTGGG | 4260 |
| TGCGCGCCAC | GTGCAGCTGG | GCAGAGGCGA | GATTCTCTCC | CGGGATCACG | AACTGTTTTA | 4320 |
| ACGGGCCGCT | CTCGGCCATA | TTGCGGTCGA | TAAGCCGCTC | CAGGGCGGTG | ATCTCTCTTT | 4380 |
| CGCCGATCGT | CTGGCTCAGG | CGGGTCAGGC | CCCGCGCATC | GCTGGCCAGT | TCAGCCCCCA | 4440 |
| GCACGAACAG | CGTCTGCTGA | ATATGGTGCA | GGCTTTCCCG | CAGCCCGGCG | TCGCGGGTCG | 4500 |
| TGGCGTAGCA | GACGCCCAGC | TGGGATATCA | GTTTCATCGAC | GGTGCCGTAG | GCCTCGACGC | 4560 |
| GAATATGGTC | TTTCTCGATG | CGGCTGCCGC | CGTACAGGGC | GGTGGTGCCT | TTATCCCCGG | 4620 |
| TGCGGGTATA | GATACGATAC | ATTCAGTTTC | TCTCACTTAA | CGGCAGGACT | TTAACCAGCT | 4680 |
| GCCCGGCGTT | GGCGCCGAGC | GTACGCAGTT | GATCGTCGCT | ATCGGTGACG | TGTCCGGTAG | 4740 |
| CCAGCGGCGC | GTCCGCCGGC | AGCTGGGCAT | GAGTGAGGGC | TATCTCGCCG | GACGCGCTGA | 4800 |
| GCCCATACC | CACCCGAGG | GGCGAGCTTC | TGGCCGCCAG | GGCGCCGAGC | GCAGCGGCGT | 4860 |

| | | | | | | |
|------------|------------|-------------|-------------|-------------|------------|------|
| CACCGCCTCC | GTCATAGGTT | ATGGTCTGGC | AGGGGACCCC | CTGCTCCTCC | AGCCCCCAGC | 4920 |
| ACAGCTCATT | GATGGCGCCG | GCATGGTGCC | CGCGCGGATC | GTAAAAACAGG | CGTACGCCTG | 4980 |
| GCGGTGAAAG | CGACATGACG | GTCCCCCTCGT | TAACACTCAG | AATGCCTGGC | GGAAAATCGC | 5040 |
| GGCAATCTCC | TGCTCGTTGC | CTTTACGCGG | GTTTCGAGAAC | GCATTGCCGT | CTTTTAGAGC | 5100 |
| CATCTCCGCC | ATGTAGGGGA | AGTCGGCCTC | TTTTACCCCC | AGATCGCGCA | GATGCTGCGG | 5160 |
| AATACCGATA | TCCATCGACA | GACGCGTGAT | AGCGGCGATG | GCTTTTTCCG | CCGCGTCGAG | 5220 |
| AGTGGACAGT | CCGGTGATAT | TTTCGCCCCAT | CAGTTCAGCG | ATATCGGCGA | ATTTCTCCGG | 5280 |
| GTTGGCGATC | AGGTTGTAGC | GCGCCACATG | CGGCAGCAGG | ACAGCGTTGG | CCACGCCGTG | 5340 |
| CGGCATGTCT | TACAGGCCGC | CCAGCTGGTG | CGCCATGGCG | TGCACGTAGC | CGAGGTTGGC | 5400 |
| GTTATTGAAA | GCCATCCCCG | CCAGCAGAGA | AGCATAGGCC | ATGTTTTCCC | GCGCCTGCAG | 5460 |
| ATTGCTGCCG | AGGGCCACGG | CCTGGCGCAG | GTTGCGGGCG | ATGAGGCGGA | TCGCCTGCAT | 5520 |
| GGCGGCGGCG | TCCGTCACCG | GGTTAGCGTC | TTTGGAGATA | TAGGCCTCTA | CGGCGTGGGT | 5580 |
| CAGGGCATCC | ATCCCGGTCT | CCGCGGTCAG | GGCGGCCGGT | TTACCGATCA | TCAGCAGTGG | 5640 |
| ATCGTTGATA | GAGACCGACG | GCAGTTTGCG | CCAGCTGACG | ATCACAAACT | TCACTTTGGT | 5700 |
| TTCGGTGTTG | GTCAGGACGC | AGTGGCGGGT | GACCTCGCTG | GCGGTGCCGG | CGGTGGTATT | 5760 |
| GACCGCGACG | ATAGGCGGCA | GCGGGTTGGT | CAGGGTCTCG | ATTCCGGCAT | ACTGGTACAG | 5820 |
| ATCGCCCTCA | TGGGTGGCGG | CGATGCCGAT | GCCTTTGCCG | CAATCGTGCG | GGCTGCCGCC | 5880 |
| GCCCACGGTG | ACGATGATGT | CGCACTGTTT | GCGGCGAAAC | ACGGCGAGGC | CGTCGCGCAC | 5940 |
| GTTGGTGTCT | TTCGGGTTCT | GCTCGACGCC | GTCAAAGATC | GCCACCTCGA | TCCCGGCCTC | 6000 |
| CCGCAGATAA | TGCAGGGTTT | TGTCCACCGC | GCCATCTTTA | ATTGCCCCGA | GGCCTTTGTC | 6060 |
| GGTGACCAGC | AGGGCTTTTT | TCCCCCCCAG | CAGCTGGCAG | CGTTCCGCCA | CTACGGAAAT | 6120 |
| GGCGTTGGGG | CCAAAAAAGT | TAACGTTTGG | CACCAGATAA | TCAAACATAC | GATAGCTCAT | 6180 |
| AATATACCTT | CTCGCTTCAG | GTTATAATGC | GGAAAAACAA | TCCAGGGCGC | ACTGGGCTAA | 6240 |
| TAATTGATCC | TGCTCGACCG | TACCGCCGCT | AACGCCGACG | GCGCCAATTA | CCTGCTCATT | 6300 |
| AAAAATAACT | GGCAGGCCGC | CGCCAAAAAT | AATAATTGCG | TGTTGGTTGG | TTAGCTGCAG | 6360 |
| ACCGTACAGA | GATTGTCCTG | GCTGGACCGC | TGACGTAATT | TCATGGGTAC | CTTGCTTCAG | 6420 |
| GCTGCAGGCG | CTCCAGGCTT | TATTCAGGGA | AATATCGCAG | CTGGAGACGA | AGGCCTCGTC | 6480 |
| CATCCGCTGG | ATAAGCAGCG | TGTTGCCTCC | GCGGTCAACT | ACGGAAAAACA | CCACCGCCAC | 6540 |
| GTTGATCTCA | GTGGCTTTTT | TTTCCACCGC | CGCCGCCATT | TGCTGGGCGG | CGGCCAGGGT | 6600 |

| | | | | | | |
|-------------|------------|-------------|------------|------------|-------------|------|
| GATTGTCTGA | ACTTGTTGGC | TCTTGTTTCAT | CATTCTCTCC | CGCACCAGGA | TAACGCTGGC | 6660 |
| GCGAATAGTC | AGTAGGGGGC | GATAGTAAAA | AACTATTACC | ATTCGGTTGG | CTTGCTTTAT | 6720 |
| TTTTGTGAGC | GTTATTTTGT | CGCCCGCCAT | GATTTAGTCA | ATAGGGTTAA | AATAGCGTCG | 6780 |
| GAAAAACGTA | ATTAAGGGCG | TTTTTTATTA | ATTGATTTAT | ATCATTGCGG | GCGATCACAT | 6840 |
| TTTTTATTTT | TGCCGCCGGA | GTAAAGTTTC | ATAGTGAAAC | TGTCGGTAGA | TTTCGTGTGC | 6900 |
| CAAATTGAAA | CGAAATTAAA | TTTATTTTTT | TCACCACTGG | CTCATTTAAA | GTTCCGCTAT | 6960 |
| TGCCGGTAAT | GGCCGGGCGG | CAACGACGCT | GGCCCGGCGT | ATTCGCTACC | GTCTGCGGAT | 7020 |
| TTCACCTTTT | GAGCCGATGA | ACAATGAAAA | GATCAAAACG | ATTTGCAGTA | CTGGCCCAGC | 7080 |
| GCCCCGTCAA | TCAGGACGGG | CTGATTGGCG | AGTGGCCTGA | AGAGGGGCTG | ATCGCCATGG | 7140 |
| ACAGCCCCTT | TGACCCGGTC | TCTTCAGTAA | AAGTGGACAA | CGGTCTGATC | GTCGAACTGG | 7200 |
| ACGGCAAACG | CCGGGACCAG | TTTGACATGA | TCGACCGATT | TATCGCCGAT | TACGCGATCA | 7260 |
| ACGTTGAGCG | CACAGAGCAG | GCAATGCGCC | TGGAGGCGGT | GGAAATAGCC | CGTATGCTGG | 7320 |
| TGGATATTCA | CGTCAGCCGG | GAGGAGATCA | TTGCCATCAC | TACCGCCATC | ACGCCGGCCA | 7380 |
| AAGCGGTCTGA | GGTGATGGCG | CAGATGAACG | TGGTGGAGAT | GATGATGGCG | CTGCAGAAGA | 7440 |
| TGCGTGCCCG | CCGGACCCCC | TCCAACCAGT | GCCACGTCAC | CAATCTCAAA | GATAATCCGG | 7500 |
| TGCAGATTGC | CGCTGACGCC | GCCGAGGCCG | GGATCCGCGG | CTTCTCAGAA | CAGGAGACCA | 7560 |
| CGGTCGGTAT | CGCGCGCTAC | GCGCCGTTTA | ACGCCCTGGC | GCTGTTGGTC | GGTTTCGAGT | 7620 |
| GCGGCCGCCC | CGGCGTGTTG | ACGCAGTGCT | CGGTGGAAGA | GGCCACCGAG | CTGGAGCTGG | 7680 |
| GCATGCGTGG | CTTAACCAGC | TACGCCGAGA | CGGTGTCGGT | CTACGGCACC | GAAGCGGTAT | 7740 |
| TTACCGACGG | CGATGATACG | CCGTGGTCAA | AGGCGTTCCT | CGCCTCGGCC | TACGCCTCCC | 7800 |
| GCGGGTTGAA | AATGCGCTAC | ACCTCCGGCA | CCGGATCCGA | AGCGCTGATG | GGCTATTTCGG | 7860 |
| AGAGCAAGTC | GATGCTCTAC | CTCGAATCGC | GCTGCATCTT | CATTACTAAA | GGCGCCGGGG | 7920 |
| TTCAGGGACT | GCAAAACGGC | GCGGTGAGCT | GTATCGGCAT | GACCGGCGCT | GTGCCGTCGG | 7980 |
| GCATTCCGGC | GGTGCTGGCG | GAAAACCTGA | TCGCCTCTAT | GCTCGACCTC | GAAGTGGCGT | 8040 |
| CCGCCAACGA | CCAGACTTTC | TCCCACTCGG | ATATTGCGCG | CACCGCGCGC | ACCCTGATGC | 8100 |
| AGATGCTGCC | GGGCACCGAC | TTTATTTTCT | CCGGCTACAG | CGCGGTGCCG | AACTACGACA | 8160 |
| ACATGTTTCG | CGGCTCGAAC | TTTCGATGCGG | AAGATTTTGA | TGATTACAAC | ATCCTGCAGC | 8220 |
| GTGACCTGAT | GGTTGACGGC | GGCCTGCGTC | CGGTGACCGA | GGCGGAAACC | ATTGCCATTTC | 8280 |
| GCCAGAAAGC | GGCGCGGGCG | ATCCAGGCGG | TTTTCCGCGA | GCTGGGGCTG | CCGCCAATCG | 8340 |

| | | | | | | |
|------------|-------------|------------|-------------|------------|-------------|-------|
| CCGACGAGGA | GGTGGAGGCC | GCCACCTACG | CGCACGGCAG | CAACGAGATG | CCGCCGCGTA | 8400 |
| ACGTGGTGGA | GGATCTGAGT | GCGGTGGAAG | AGATGATGAA | GCGCAACATC | ACCGGCCTCG | 8460 |
| ATATTGTCGG | CGCGCTGAGC | CGCAGCGGCT | TTGAGGATAT | CGCCAGCAAT | ATTCTCAATA | 8520 |
| TGCTGCGCCA | GCGGGTCACC | GGCGATTACC | TGCAGACCTC | GGCCATTCTC | GATCGGCAGT | 8580 |
| TCGAGGTGGT | GAGTGCGGTC | AACGACATCA | ATGACTATCA | GGGGCCGGGC | ACCGGCTATC | 8640 |
| GCATCTCTGC | CGAACGCTGG | GCGGAGATCA | AAAATATTCC | GGGCGTGGTT | CAGCCCGACA | 8700 |
| CCATTGAATA | AGGCGGTATT | CCTGTGCAAC | AGACAACCCA | AATTCAGCCC | TCTTTTACCC | 8760 |
| TGAAAACCCG | CGAGGGCGGG | GTAGCTTCTG | CCGATGAACG | CGCCGATGAA | GTGGTGATCG | 8820 |
| GCGTCGGCCC | TGCCTTCGAT | AAACACCAGC | ATCACACTCT | GATCGATATG | CCCCATGGCG | 8880 |
| CGATCCTCAA | AGAGCTGATT | GCCGGGGTGG | AAGAAGAGGG | GCTTCACGCC | CGGGTGGTGTC | 8940 |
| GCATTCTGCG | CACGTCCGAC | GTCTCCTTTA | TGGCCTGGGA | TGCGGCCAAC | CTGAGCGGCT | 9000 |
| CGGGGATCGG | CATCGGTATC | CAGTCGAAGG | GGACCACGGT | CATCCATCAG | CGCGATCTGC | 9060 |
| TGCCGCTCAG | CAACCTGGAG | CTGTTCTCCC | AGGCGCCGCT | GCTGACGCTG | GAGACCTACC | 9120 |
| GGCAGATTGG | CAAAAACGCT | GCGCGCTATG | CGCGCAAAGA | GTCACCTTCG | CCGGTGCCGG | 9180 |
| TGGTGAACGA | TCAGATGGTG | CGGCCGAAAT | TTATGGCCAA | AGCCGCGCTA | TTTCATATCA | 9240 |
| AAGAGACCAA | ACATGTGGTG | CAGGACGCCG | AGCCCGTCAC | CCTGCACATC | GACTTAGTAA | 9300 |
| GGGAGTGACC | ATGAGCGAGA | AAACCATGCG | CGTGCAGGAT | TATCCGTTAG | CCACCCGCTG | 9360 |
| CCCGGAGCAT | ATCCTGACGC | CTACCGGCAA | ACCATTGACC | GATATTACCC | TCGAGAAGGT | 9420 |
| GCTCTCTGGC | GAGGTGGGCC | CGCAGGATGT | GCGGATCTCC | CGCCAGACCC | TTGAGTACCA | 9480 |
| GGCGCAGATT | GCCGAGCAGA | TGCAGCGCCA | TGCGGTGGCG | CGCAATTTCC | GCCGCGCGGC | 9540 |
| GGAGCTTATC | GCCATTCTCTG | ACGAGCGCAT | TCTGGCTATC | TATAACGCGC | TGCGCCCGTT | 9600 |
| CCGCTCCTCG | CAGGCGGAGC | TGCTGGCGAT | CGCCGACGAG | CTGGAGCACA | CCTGGCATGC | 9660 |
| GACAGTGAAT | GCCGCCTTTG | TCCGGGAGTC | GGCGGAAGTG | TATCAGCAGC | GGCATAAGCT | 9720 |
| GCGTAAAGGA | AGCTAAGCGG | AGGTCAGCAT | GCCGTTAATA | GCCGGGATTG | ATATCGGCAA | 9780 |
| CGCCACCACC | GAGGTGGCGC | TGGCGTCCGA | CTACCCGCAG | GCGAGGGCGT | TTGTTGCCAG | 9840 |
| CGGGATCGTC | GCGACGACGG | GCATGAAAGG | GACGCGGGAC | AATATCGCCG | GGACCCTCGC | 9900 |
| CGCGCTGGAG | CAGGCCCTGG | CGAAAACACC | GTGGTTCGATG | AGCGATGTCT | CTCGCATCTA | 9960 |
| TCTTAACGAA | GCCGCGCCGG | TGATTGGCGA | TGTGGCGATG | GAGACCATCA | CCGAGACCAT | 10020 |
| TATCACCGAA | TCGACCATGA | TCGGTCATAA | CCCGCAGACG | CCGGGCGGGG | TGGGCGTTGG | 10080 |

| | | | | | | |
|-------------|------------|------------|-------------|------------|------------|-------|
| CGTGGGGACG | ACTATCGCCC | TCGGGCGGCT | GGCGACGCTG | CCGGCGGCGC | AGTATGCCGA | 10140 |
| GGGGTGGATC | GTACTGATTG | ACGACGCCGT | CGATTTCCTT | GACGCCGTGT | GGTGGCTCAA | 10200 |
| TGAGGCGCTC | GACCGGGGGA | TCAACGTGGT | GGCGGCGATC | CTCAAAAAGG | ACGACGGCGT | 10260 |
| GCTGGTGAAC | AACCGCCTGC | GTAAAACCCT | GCCGGTGGTG | GATGAAGTGA | CGCTGCTGGA | 10320 |
| GCAGGTCCCC | GAGGGGGTAA | TGGCGGCGGT | GGAAGTGGCC | GCGCCGGGCC | AGGTGGTGCG | 10380 |
| GATCCTGTCG | AATCCCTACG | GGATCGCCAC | CTTCTTCGGG | CTAAGCCCGG | AAGAGACCCA | 10440 |
| GGCCATCGTC | CCCATCGCCC | GCGCCCTGAT | TGGCAACCGT | TCCGCGGTGG | TGCTCAAGAC | 10500 |
| CCCGCAGGGG | GATGTGCAGT | CGCGGGTGAT | CCCGGCGGGC | AACCTCTACA | TTAGCGGCGA | 10560 |
| AAAGCGCCGC | GGAGAGGCCG | ATGTCGCCGA | GGGCGCGGAA | GCCATCATGC | AGGCGATGAG | 10620 |
| CGCCTGCGCT | CCGGTACGCG | ACATCCGCGG | CGAACC GGCG | ACCCACGCCG | GCGGCATGCT | 10680 |
| TGAGCGGGTG | CGCAAGGTAA | TGGCGTCCCT | GACCGGCCAT | GAGATGAGCG | CGATATACAT | 10740 |
| CCAGGATCTG | CTGGCGGTGG | ATACGTTTAT | TCCGCGCAAG | GTGCAGGGCG | GGATGGCCGG | 10800 |
| CGAGTGCGCC | ATGGAGAATG | CCGTCGGGAT | GGCGGCGATG | GTGAAAGCGG | ATCGTCTGCA | 10860 |
| AATGCAGGTT | ATCGCCCGCG | AACTGAGCGC | CCGACTGCAG | ACCGAGGTGG | TGGTGGGCGG | 10920 |
| CGTGGAGGCC | AACATGGCCA | TCGCCGGGGC | GTAAACCACT | CCCGGCTGTG | CGGCGCCGCT | 10980 |
| GGCGATCCTC | GACCTCGGCG | CCGGCTCGAC | GGATGCGGCG | ATCGTCAACG | CGGAGGGGCA | 11040 |
| GATAACGGCG | GTCCATCTCG | CCGGGGCGGG | GAATATGGTC | AGCCTGTTGA | TTAAAACCGA | 11100 |
| GCTGGGCCTC | GAGGATCTTT | CGCTGGCGGA | AGCGATAAAA | AAATACCCGC | TGGCCAAAGT | 11160 |
| GGAAAGCCTG | TTCAGTATTC | GTCACGAGAA | TGGCGCGGTG | GAGTTCTTTC | GGGAAGCCCT | 11220 |
| CAGCCCGGCG | GTGTTGCGCA | AAGTGGTGTA | CATCAAGGAG | GGCGAACTGG | TGCCGATCGA | 11280 |
| TAACGCCAGC | CCGCTGGAAA | AAATTGCTCT | CGTGCGCCCG | CAGGCGAAAG | AGAAAGTGTT | 11340 |
| TGTCACCAAC | TGCCTGCGCG | CGCTGCGCCA | GGTCTCACCC | GGCGGTTCCA | TTGCGGATAT | 11400 |
| CGCCTTTTGTG | GTGCTGGTGG | GCGGCTCATC | GCTGGACTTT | GAGATCCCGC | AGCTTATCAC | 11460 |
| GGAAGCCTTG | TCGCACTATG | GCGTGGTCGC | CGGGCAGGGC | AATATTCGGG | GAACAGAAGG | 11520 |
| GCCGCGCAAT | GCGGTGCGCA | CCGGGCTGCT | ACTGGCCGGT | CAGGCGAATT | AAACGGGCGC | 11580 |
| TCGCGCCAGC | CTCTCTCTTT | AACGTGCTAT | TTCAGGATGC | CGATAATGAA | CCAGACTTCT | 11640 |
| ACCTTAACCG | GGCAGTGCGT | GGCCGAGTTT | CTTGGCACCG | GATTGCTCAT | TTTCTTCGGC | 11700 |
| GCGGGCTGCG | TCGCTGCGCT | GCGGGTCGCC | GGGGCCAGCT | TTGGTCAGTG | GGAGATCAGT | 11760 |
| ATTATCTGGG | GCCTTGCGGT | CGCCATGGCC | ATCTACCTGA | CGGCCGGTGT | CTCCGGCGCG | 11820 |

CACCTAAATC CGGCGGTGAC CATTGCCCTG TGGCTGTTCTG CCTGTTTTGA ACGCCGCAAG 11880
 GTGCTGCCGT TTATTGTTGC CCAGACGGCC GGGGCCTTCT GCGCCGCCGC GCTGGTGTAT 11940
 GGGCTCTATC GCCAGCTGTT TCTCGATCTT GAACAGAGTC AGCATATCGT GCGCGGCACT 12000
 GCCGCCAGTC TTAACCTGGC CGGGGTCTTT TCCACGTACC CGCATCCACA TATCACTTTT 12060
 ATACAAGCGT TTGCCGTGGA GACCACCATC ACGGCAATCC TGATGGCGAT GATCATGGCC 12120
 CTGACCGACG ACGGCAACGG AATTC 12145

(2) INFORMATION FOR SEQ ID NO:2:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 30 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

GGGAATTCAT GAAAAGATCA AAACGATTTG 30

(2) INFORMATION FOR SEQ ID NO:3:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 29 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:

GCGAATTCTT ATTCAATGGT GTCGGGCTG 29

(2) INFORMATION FOR SEQ ID NO:4:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 30 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:

GCGAATTCAT GCAACAGACA ACCCAAATTC 30

(2) INFORMATION FOR SEQ ID NO: 5:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 25 base pairs

- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:5:

GCGAATTCAC TCCCTTACTA AGTCG

25

(2) INFORMATION FOR SEQ ID NO: 6:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 27 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:6:

GCGAATTCAT GAGCTATCGT ATGTTTG

27

(2) INFORMATION FOR SEQ ID NO:7:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 28 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:7:

GCGAATTCAG AATGCCTGGC GGAAAATC

28

(2) INFORMATION FOR SEQ ID NO:8:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 28 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:8:

GGGAATTCAT GAGCGAGAAA ACCATGCG

28

(2) INFORMATION FOR SEQ ID NO:9:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 27 base pairs
 - (B) TYPE: nucleic acid

(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:9:

GCGAATTCTT AGCTTCCTTT ACGCAGC

27

(2) INFORMATION FOR SEQ ID NO:10:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 94 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:10:

AGCTTAGGAG TCTAGAATAT TGAGCTCGAA TTCCCGGGCA TGCGGTACCG GATCCAGAAA

60

AAAGCCCGCA CCTGACAGTG CGGGCTTTTT TTTT

94

(2) INFORMATION FOR SEQ ID NO:11:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 33 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:11:

GGCCAAGCTT AAGGAGGTTA ATTAAATGAA AAG

33

(2) INFORMATION FOR SEQ ID NO:12:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 26 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:12:

GCTCTAGATT ATTCAATGGT GTCGGG

26

(2) INFORMATION FOR SEQ ID NO:13:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 42 base pairs
(B) TYPE: nucleic acid

(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:13:

GCGCCGTCTA GAATTATGAG CTATCGTATG TTTGATTATC TG

42

(2) * INFORMATION FOR SEQ ID NO:14:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 36 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:14:

TCTGATACGG GATCCTCAGA ATGCCTGGCG GAAAAT

36

(2) INFORMATION FOR SEQ ID NO:15:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 181 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:15:

CGATCTGTGC TGTTTGCCAC GGTATGCAGC ACCAGCGCGA GATTATGGGC TCGCACGCTC 60
GACTGTCGGA CGGGGGCACT GGAACGAGAA GTCAGGCGAG CCGTCACGCC CTTGACAATG 120
CCACATCCTG AGCAAATAAT TCAACCACTA AACAAATCAA CCGCGTTTCC CGGAGGTAAC 180
C 181

(2) INFORMATION FOR SEQ ID NO:16:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 149 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:16:

CGATCTGTGC TGTTTGCCAC GGTATGCAGC ACCAGCGCGA GATTATGGGC TCGCACGCTC 60
GACTGTCGGA CGGGGGCACT GGAACATGCC ACATCCTGAG CAAATAATTC AACCCTAAA 120

CAAATCAACC GCGTTTCCCG GAGGTAACC

149

(2) INFORMATION FOR SEQ ID NO:17:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 33 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:17:

GGAATTCAC T AGTCGATCTG TGCTGTTTGC CAC

33

(2) INFORMATION FOR SEQ ID NO:18:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 33 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:18:

GGGGAAGCTT GGTTACCTCC GGGAAACGCG GTT

33

(2) INFORMATION FOR SEQ ID NO:19:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 16 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:19:

TCGACCACAA GGAGGA

16

(2) INFORMATION FOR SEQ ID NO:20:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 16 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:20:

CTAGTCCTCC TTGTGG

16

002250" 05957560

(2) INFORMATION FOR SEQ ID NO:21:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 45 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:21:

ACTGGCCGTC GTTTACTCG AGTCGTGACT GGGAAAACCC TGGCG

45

(2) INFORMATION FOR SEQ ID NO:22:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 14 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:22:

AATTCAAAGG AGGT

14

(2) INFORMATION FOR SEQ ID NO:23:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 14 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:23:

CTAGACCTCC TTTG

14

(2) INFORMATION FOR SEQ ID NO:24:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 19 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:24:

AGCTTGTCGA CCATGAAAA

19

002250" 052200

(2) INFORMATION FOR SEQ ID NO:25:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 19 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: DNA (genomic)
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:25:

GATCTTTTCA TGGTCGACA

19

(2) INFORMATION FOR SEQ ID NO:26:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 13 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: DNA (genomic)
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:26:

TCGACCAGGA GGA

13

(2) INFORMATION FOR SEQ ID NO:27:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 13 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: DNA (genomic)
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:27:

CTAGTCCTCC TGG

13

(2) INFORMATION FOR SEQ ID NO:28:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 18 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: DNA (genomic)
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:28:

TCGACGAATT CAGGAGGA

18

(2) INFORMATION FOR SEQ ID NO:29:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 18 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: DNA (genomic)
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:29:

CTAGTCCTCC TGAATTCG

18

(2) INFORMATION FOR SEQ ID NO:30:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 23 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: DNA (genomic)
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:30:

ATGTACAAGA TCCTGATCGC CGA

23

(2) INFORMATION FOR SEQ ID NO:31:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 22 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: DNA (genomic)
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:31:

TCAGCGGCGC AGGTAGGCGG CG

22

(2) INFORMATION FOR SEQ ID NO:32:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 31 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: DNA (genomic)
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:32:

ATGACCAAGG GCCGGATCCG TCGACCTGCA G

31

(2) INFORMATION FOR SEQ ID NO:33:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 32 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:33:

CTACCCTTGG CCCCGGATCC GTCGACCTGC AG

32

(2) INFORMATION FOR SEQ ID NO:34:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 23 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:34:

CACGGCCTGG CGCAGGTTGC GGG

23

(2) INFORMATION FOR SEQ ID NO:35:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 21 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:35:

GGCAGCCCGC ACGATTGCGG C

21

(2) INFORMATION FOR SEQ ID NO:36:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 21 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:36:

GCGGAAAACC GCCTGGATCG C

21

(2) INFORMATION FOR SEQ ID NO:37:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 21 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: DNA (genomic)
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:37:

GGG TTCAGG ACTGCAAAAC G

21

(2) INFORMATION FOR SEQ ID NO:38:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 35 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: DNA (genomic)
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:38:

GGAATTCAGA TCTCAGCAAT GAAAAGATCA AAACG

35

(2) INFORMATION FOR SEQ ID NO:39:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 34 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: DNA (genomic)
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:39:

GGAATTCAGA TCTCAGCAAT GCAACAGACA ACCC

34

(2) INFORMATION FOR SEQ ID NO:40:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 28 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: DNA (genomic)
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:40:

GCTCTAGATC ACTCCCCTTA CTAAGTCG

28

[illegible]

(A) LENGTH: 37 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:41:

37

(A) LENGTH: 27 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:42:

27

(A) LENGTH: 38 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:43:

38

(A) LENGTH: 24 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:44:

24

(2) INFORMATION FOR SEQ ID NO:45:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 35 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:45:

GGAATTCAGA TCTAGCAATG CCGTTAATAG CCGGG

35

(2) INFORMATION FOR SEQ ID NO:46:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 26 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:46:

GCTCTAGATT AATTCGCCTG ACCGGC

26